

## Marked-Up Version of Substitute Specification

~~Description~~ TITLE OF THE INVENTION

~~Virtual assistant, which outputs audible information to a user of a data terminal by means of at least two electroacoustic converters, and method for presenting audible information of a virtual assistant~~

### VIRTUAL ASSISTANT AND METHOD FOR PROVIDING AUDIBLE INFORMATION TO A USER

#### BACKGROUND OF THE INVENTION

The invention relates to a virtual assistant, which outputs audible information to a user of a data terminal by means of at least two electroacoustic converters, and a method for presenting audible information of a virtual assistant for a user of a data terminal.

When using PC application programs, it is generally known that the user can make use of a virtual assistant, ~~that~~. A virtual assistant is to say a computer-based help (program) program that supports the user when carrying out the steps necessary to perform a task, or on the computer. The virtual assistant may also be invoked when the user wishes needs further explanations about the capabilities of the PC application program. Secondly, The virtual assistant may also direct the user's attention is drawn to any incorrect inputs and the virtual assistant makes input to any input mistakes the user makes and may make suggestions to the user. The information provided by the virtual assistant is presented to the user ~~optically~~ visually, that is to say by means of a display unit.

In principle, ~~this function~~ the functions of a virtual assistant which ~~is~~ are helpful to ~~the~~ a user can also be applied to mobile data terminals such as mobile phones or handheld terminals that are known as Personal Digital Assistants (PDAs). In this case, however, ~~it is disadvantageous for the user that the extensive information presented by the user of visual data by a traditional virtual assistant must be displayed on a~~ is a disadvantage due to the small display unit of the mobile data terminal.

Moreover, the extensive amount of information of presented visually by a virtual assistant that is presented optically is difficult for the user of a handheld data

terminal to process ~~whenever the user needs to~~ in situations where the user must concentrate on other ~~optically~~visually presented information presented in the same vicinity or on acoustic information ~~of a~~simultaneously presented such as an ongoing conversation ~~partner at the same time~~with an associate. In this case it is expedient to provide the information presented by the virtual assistant ~~of a data terminal for~~to the data terminal user by means of an acoustic presentation. In this way, the data terminal user can ~~better~~more easily process the acoustically presented information ~~and additional~~along with the other information ~~optically presented~~being simultaneously presented either visually or accoustically.

~~On the~~In other ~~hand~~applications, data terminals ~~or methods~~ are known in ~~which~~employed where additional information is acoustically presented to the user ~~of the data terminal or of the method~~accoustically. For instance, an audio assistant in a ticket machine ~~guides them~~may be used to guide a user of the ticket machine through the respective operating programs of the ticket machine ~~by means of acoustic information~~. Since ~~these~~ However, such ticket machines and like devices are often sited in ~~a loud environment, it is~~noisy environments. It is often difficult for ~~the user~~users of the ticket machine to ~~follow~~hear the acoustic information output by the audio assistant of the ticket machine. ~~It and follow the instructions being presented~~.

An additional complicating factor in presenting acoustic information is that it is even more difficult to follow acoustic information that is simultaneously acting on a user from two different signal sources. So-called binaural technology has been the subject of research for some time now. For example, an introduction to binaural technology is described under the title: "An introduction to binaural technology" by J. Blauert (1996) in Binaural and Spatial Hearing in Real and Virtual Environments, edited by R. Gilkey & T. Anderson, pages 593-609, Lawrence Erlbaum, USA-Hilldale NJ.

With the aid of binaural technology, ~~using~~ signal processing of the sound information can be employed to give, the listener ~~can assign~~the sense that the sound-generating source is assigned to any ~~positions of~~position within the surrounding space. ~~The position~~Though the relative positions of the listener ~~here,~~

- and of the electroacoustic converters outputting the acoustic information
- ~~respectively, remains~~remain spatially fixed. ~~By means of suitable signal processing of the sound information, it is then possible for example, it is possible~~ to awaken in the listener the subjective impression that the sound-generating source is turning around him, ~~or is coming~~moving toward him, ~~or is moving away from him, or~~ changing in some other way. By signal processing of the sound information, the sound-generating source can ~~therefore be spatially positioned anywhere. It is therefore the object of the present invention to develop a technical solution for the user of a data terminal, in which the acoustic information output by the virtual assistant of the data terminal can be better separated, in terms of the user's perception, from other sound sources that are likewise acting on the data terminal user~~ be positioned anywhere in space, yet give the impression to the user that it is located elsewhere.
- The ~~object is achieved on the basis of the virtual assistant defined in the preamble of claim 1 by the features set out in the characterizing part of claim 1, and on the basis of the method defined in the preamble of claim 9 by the features set out in the characterizing part of claim 9. Advantageous refinements of the invention are set out in the subclaims.~~

## SUMMARY OF THE INVENTION

According to the invention, a virtual assistant which outputs audible information to a data terminal user by means of at least two electroacoustic converters can be spatially positioned by the user in order to achieve a better spatially acoustic separation between the information output by means of the electroacoustic converters and additional information output by at least one ~~further~~other sound source.

~~One~~An advantage of the invention is the ~~utilization of the spatial positioning of sound sources by means of~~ is that signal processing of the sound information of the virtual assistant ~~may utilize the spatial positioning of the sound sources of~~ relative to the data terminal, ~~or its locating by the data terminal user respectively.~~

For the data terminal user, said sound information of user so that the virtual assistant can be better perceived separately from ambient noises~~noise~~.

Furthermore, the sound information of the virtual assistant can be supplied to the data terminal user in a targeted manner from ~~one a~~ specific direction, while the user is simultaneously holding a conversation with ~~another conversation partner~~ someone else in the room. Here, too, it is possible to achieve a ~~good~~ satisfactory spatially acoustic separation between the sound information acting on the user from the virtual assistant and from the ~~conversation partner~~ person conversing with the user. This enables the user to receive and process both the information coming from the virtual assistant and the information coming from ~~the~~ his conversation partner. The simultaneous ~~receiving~~ reception and processing of both ~~the sets of~~ information coming from the virtual assistant and the information coming from the conversation partner ~~is however~~ is at least facilitated for the user.

A further advantage emerges when, in addition to the sound information coming from the virtual assistant and the ambient noises originating from ~~further~~ other sound sources present in the vicinity of the user, visual information is also ~~information that is additionally optically presented simultaneously acts onto~~ the data terminal user at the same time. In this case, too, the data terminal user can better receive and process the information coming from the various sources. ~~Further~~

Additional features and advantages of the present invention emerge from the description below, in which the invention is explained with reference to two exemplary embodiments. are described in, and will be apparent from, the following Detailed Description of the Invention and the figures.

## DETAILED DESCRIPTION OF THE INVENTION

In ~~the~~ a first ~~exemplary~~ embodiment, a pedestrian is situated in road traffic. The pedestrian is laden with heavy shopping bags. The pedestrian would like to conduct a phone call using his data terminal in the form of a mobile phone. The mobile phone is switched on, but is stowed away in one of his shopping bags and therefore cannot be readily located. The pedestrian is wearing a light headphones and microphone set however. Integrated in the headphones and microphone set are

two electroacoustic converters for outputting sound information. Like the mobile phone, the headphones and microphone set is connected to a radio module, for example to a Bluetooth radio module, for short-range data exchange between the headphones and microphone set and the mobile phone.

5           The pedestrian, user of the headphones and microphone set and of the mobile phone respectively, activates the headphones and microphone set and thus enables data exchange between the headphones and microphone set and the mobile phone. The user speaks the word "DIAL" into the headphones and microphone set, whereupon the virtual assistant of the mobile phone responds with "PLEASE SAY  
10   THE NAME". The user says the name of the person he wishes to call. Since the user is moving in an environment with a high noise level, the mobile phone does not recognize the name of the person to be called with sufficient accuracy. The mobile phone processes the name entered by the user and compares it with names stored in the internal phone directory of the mobile phone. The mobile phone  
15   recognizes the name spoken as "SCHMITZER" or "SCHNITZLER". Output of the two names to the display unit of the mobile phone and the subsequent request to the user to select one of these names is of no use to the user. ~~This is because, as already mentioned, the user's mobile phone is hidden in one of the pedestrian's shopping bags in a place that is difficult to access. On the other hand~~ However, the mobile  
20   phone has recognized the ~~operation of the mobile phone~~ request by the user via the headphones and microphone set, so the mobile phone instructs the virtual assistant ~~of the mobile phone to output all similarly sounding names to the user by means of the virtual assistant using~~ the headphones and microphone set. For example, the user hears the following words of his virtual assistant via the headphones and  
25   microphone set: "THE NAME WAS NOT CLEARLY RECOGNIZED". "PLEASE SELECT ONE OF THE FOLLOWING OPTIONS". "SCHMITZER" or after a brief pause "SCHNITZLER".

          Despite the loud ambient noises, the user recognizes both the options offered by the virtual assistant because binaural technology is used during the  
30   output of the sound information of the virtual assistant of the mobile phone by means of the electroacoustic converters. The binaural technology enables the

targeted signal processing of the sound information ~~is~~ output by the mobile phone. When the sound information is played back by the virtual assistant using the headphones and microphone set, the mobile phone user can perceive a clear local attribution of the sound information output by the virtual assistant. In accordance  
5 with a user preset, in the mobile phone, the sound information is processed using signal technology in such a way that the mobile phone user locates the sound information presented by the virtual assistant as if it were coming from the vicinity of the head. The sound information is "whispered" into the user's ear over his shoulder from behind.

10 The position of the virtual assistant, or the position from which the sound information output by the virtual assistant is perceived respectively, can be changed as desired by the mobile phone user, for example by means of an electromechanical input device as is well known per-se in the art.

The electromechanical input device ~~is~~ may be for example a ball ~~in a~~  
15 socket. ~~The input device, where the~~ rotations of the ball produced by the user are detected by sensors. Alternatively, the positioning of the virtual assistant ~~is~~ may be performed ~~in a manner known per-se~~ by means of voice commands or by means of inputs on a touch-sensitive display unit of the mobile phone.

If the mobile phone has a head position sensor which detects the head movements  
20 of the mobile phone user, for example using a rotational rate sensor or a magnetic field sensor, it is furthermore possible for the selected position of the virtual assistant to be retained even if the head movements are taken into account during the signal processing of the sound information.

By means of the preset positioning of the virtual assistant, or ~~respectively~~  
25 the ability of the user to change its position as desired, the user can both operate the mobile phone in a simple manner using voice commands to establish an outgoing connection as well as attentively perceive ambient noises, such as loud calls or the sounding of horns etc.

To finish the selection of the names "SCHMITZER" or "SCHNITZLER"  
30 presented by the virtual assistant in order to establish an outgoing connection, the user responds to the name "SCHMITZER" by speaking a "NO" into the headphones

and microphone set and by responding "YES" for the name "SCHNITZLER". The mobile phone recognizes the name "SCHNITZLER" and establishes an outgoing call.

In ~~the~~ a second exemplary embodiment, a teleconferencing situation is described. Taking part in the teleconference ~~are~~ is established among a plurality of people ~~who for the most part, many of whom~~ speak and understand different languages. The ~~people~~ participants in the teleconference are situated in each case at individual tables spread throughout ~~the~~ a teleconferencing room, ~~with each.~~ Each person ~~having~~ has their own display. If one participant starts to speak, ~~then the~~ a data terminal in the form of a teleconferencing system displays ~~this~~ the participant on a large screen on a side wall of the teleconferencing room, so that the other participants can ~~also observe~~ the facial expressions and gestures of ~~this~~ the participant who is speaking.

Secondly, ~~his~~ the speaker's speech is output via electroacoustic converters in the form of loudspeakers which are connected to the teleconferencing system.

At the same time, the ~~contributions of the speaking participant are~~ speaker's speech is simultaneously interpreted into the languages of the other participants, ~~and the translation is.~~ The translations are made available to the participants in the form of sound information via ~~a~~ headphones and microphone ~~set~~ sets in which two electroacoustic converters for outputting sound information are integrated. To offer the participants the option of attentively following the speech both in the language of the participant speaking and in the language of the simultaneous interpretation, the simultaneous interpretation is output by the teleconferencing system using a virtual assistant so that the other participants can hear it. The virtual assistant can be positioned anywhere in the room by each teleconference participant by entering the respective key combinations into the teleconferencing system.

Here, too, the positioning of the virtual assistant, or the spatially acoustic perception of the sound information output by the virtual assistant by the individual participants respectively, is achieved by means of signal processing of the sound information in the teleconferencing system. The participants position the virtual assistant in such a way that the participants perceive the output of the sound

information by the virtual assistant as being transmitted over the shoulder from behind and coming from the vicinity of the head. By virtue of this positioning of the virtual assistant, a good spatially acoustic separation between the speech transmitted via loudspeakers and the simultaneous interpretation of the speech is achieved, ~~so that the~~. The participants can readily follow both the speech transmitted via loudspeakers and the simultaneous ~~interpretation, and can still~~translation while attentively ~~observe~~observing the facial expressions and gestures of the participant speaking. That is to say, the participants can ~~highly~~ attentively follow a plurality of information streams at the same time.

If one participant already knows what one of his own delegation is going to say, then said participant can have the teleconferencing system acoustically give him further information via the virtual assistant, for example about the schedule for the day, background information about the other participants, or information about the participant's hotel.

The above embodiments of the invention are merely examples given and are not exhaustive. The concept of ~~the~~spatially acoustic separation and signal processing of sound information which is output to a data terminal user via a virtual assistant and additional simultaneously audible and/or visible information which is important to the user can be applied to further examples, ~~in~~. In particular, the present invention may also be employed in cases where mobile communication terminals are employed by a user. Travel guides are cited here by way of example, wherein the travel guide explains certain exhibits of a museum to visitors in the local language of the country/museum; the visitors are able to listen ~~via their UMTS mobile phone~~ to a simultaneous ~~interpretation~~translation of the explanations of the travel guide ~~with~~on their UMTS mobile phone having good spatially acoustic separation via a virtual assistant, ~~and optionally~~. Optionally the user can attentively follow additional optical information relating to the exhibits on the display unit of their UMTS mobile phone at the same time.



### ABSTRACT OF THE DISCLOSURE

5 A virtual assistant is provided which outputs audible information to a user of a data terminal. A method is also provided for presenting audible information of a virtual assistant to the user of a data terminal. At least two electroacoustic converters are driven such that the virtual assistant can be spatially positioned by a data terminal use to achieve acoustic separation between the electroacoustic converters and an additional sound source.